Exact analytical solution of binary dynamics on networks

E. Laurence[1], J.-G. Young[1], S. Melnik[2] and L.J. Dubé[1]

[1]*Département de physique, de génie physique, et d'optique, Université Laval, Québec, Canada*
[2]*MACSI, Departement of Mathematics & Statistics, University of Limerick, Ireland*

Summary

Cascade dynamics have been extensively studied using analytical formalisms suited to a broad class of dynamical models [1]. Noteworthy examples include site and bond percolation, Watt's model of threshold dynamics, and Susceptible-Infected-Removed (SIR) disease transmission.

Using recursive equations, Ref [2] was able to solve exactly bond percolation on arbitrary graph. We extend this method to cascade dynamics.

We exaclty solve cascade dynamics on arbitrary networks.

Also, we propose two algoritms to compute the solution.
Our method has a number of advantages:

- Single computation for all dynamics.
- Exact evaluation of probability of rare events.
- Multiple results can be obtained exactly : size distribution, mean size of giant component, probability of activation of a single node...

Binary Cascade Dynamics

- Nodes can be in any one of two states (inactive, active), and irreversibly change from inactive to active, as soon as their neighborhood satisfies a predetermined condition.

Process

1. Assign a threshold to each node following a response function.
2. Activate a seed node.
3. Activate nodes whose number of active neighbors exceed their threshold. Repeat this step until no futher change is possible.

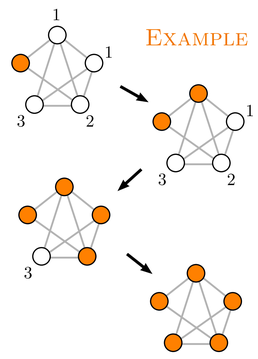
Response Function

$F(m,k)$: Probability for a node of degree k to be active when it has m active neighbors.

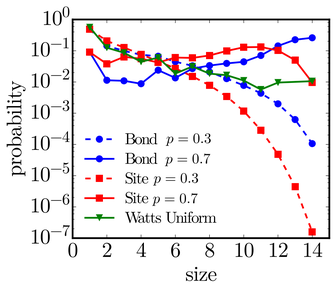
Examples of $F(m,k)$ for specific processes [1] :

Bond Percolation	Site Percolation	Watts Model
$1-(1-p)^m$	$Q_k(1-\delta_{m,0})$	$C(m/k)$

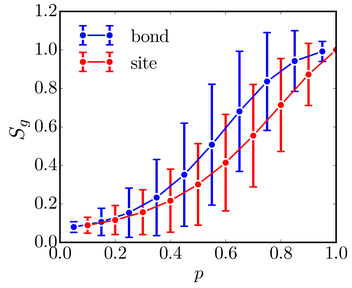
Example



Case Study IEEE 14 Bus Network



The size distributions are obtained by summing the probabilities over configurations of the same size under the same dynamical process.



The mean size of the giant component can be effectively computed from the size distribution averaged over all configurations and the inherent variance of the size distribution can be obtained exactly (error bars).

Advantages

- Single computation allows us to solve for all dynamics on a given network. One only needs to choose the appropriate response function.
- Exact probability for configurations that occur so rarely that their numerical evaluation using Monte Carlo simulations is virtually impossible.
- Exact size distribution and mean size of the giant component, with associated error bars.

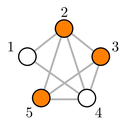
Formalism

Consider a graph composed of N nodes.

Let $\boldsymbol{l}$ be a vector of length N describing a subgraph whose components l_i is 1 if node i belongs to the subgraph.

Let $\boldsymbol{n}$ be the equivalent vector for the entire graph.

$Q(\boldsymbol{l}|\boldsymbol{n})$: Probability that every component of the subgraph $\boldsymbol{l}$ of graph $\boldsymbol{n}$ are active.



$N=5$
$\boldsymbol{l}=(0,1,1,0,1)$
$\boldsymbol{n}=(1,1,1,1,1)$

Then:

The following recursive set of equations exactly solves the cascade dynamics:

$$Q(\boldsymbol{l}|\boldsymbol{n}) = Q(\boldsymbol{l}|\boldsymbol{l}) \prod_{i=1}^{N} [1-F(m_i,k_i)]^{n_i-l_i} \equiv Q(\boldsymbol{l}|\boldsymbol{l}) \prod_{i=1}^{N} [G(m_i,k_i)]^{n_i-l_i}$$
$$Q(\boldsymbol{l}|\boldsymbol{l}) = 1 - \textstyle\sum_{\boldsymbol{u}<\boldsymbol{l}} Q(\boldsymbol{u}|\boldsymbol{l})$$

The key is to multiply the probability to activate the subgraph $\boldsymbol{l}$ by the probability that the inactive nodes have a threshold higher than their number of active neighbors.

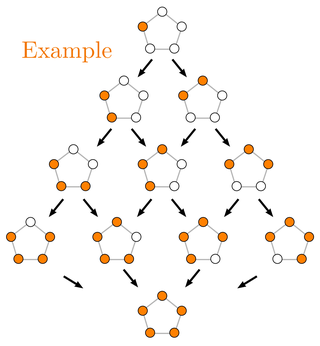
Modified Breadth-First Method

- A configuration is said possible if every active node can be reached from the seed node by following a path consisting only of active nodes.
- For sparse graphs, the majority of the configurations are impossible.

Algorithm

1. Start with the configuration where only the seed node is activated.
2. Activate a node adjacent to the active component and calculate its probability $Q(\boldsymbol{l}|\boldsymbol{n})$. Repeat this step.

Example



Maths

Using

$$Q(\boldsymbol{u}|\boldsymbol{l}) = Q(\boldsymbol{u}|\boldsymbol{n}) \left[\prod_{i=1}^{N} G(m_i(\boldsymbol{u}),k_i)^{n_i-l_i}\right]^{-1}$$

we can write

$$Q(\boldsymbol{l}|\boldsymbol{l}) = 1 - \textstyle\sum_{\boldsymbol{u}<\boldsymbol{l}} Q(\boldsymbol{u}|\boldsymbol{n}) \left[\prod_{i=1}^{N} G(m_i(\boldsymbol{u}),k_i)^{n_i-l_i}\right]^{-1}$$

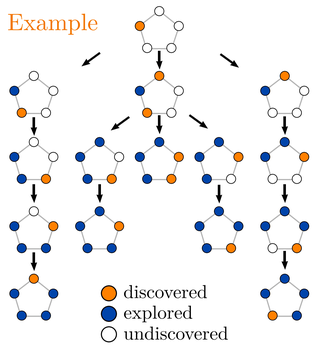
Then, the probability $Q(\boldsymbol{l}|\boldsymbol{l})$ can be expressed as the sum over smaller configurations times a coefficient.

Since the sum is over $\boldsymbol{u}<\boldsymbol{l}$, every $Q(\boldsymbol{u}|\boldsymbol{n})$ has already been computed previously by the algorithm.

Modified Depth-First Method

- Nodes are labeled as undiscovered, *discovered* and explored. Every node is initially undiscovered. A node stays in discovered state for one algorithm step; it is then marked as explored.

Example



Algorithm

1. Start with a single discovered node (seed).
2. Enumerate every configurations containing at least one undiscovered neighbor of the discovered node(s).
3. For each new configuration, mark the undiscovered neighbors as discovered, and return to step 2.
4. When there is no undiscovered neighbors left, backtrack and repeat steps 2-3.

The probability $Q(\boldsymbol{l}|\boldsymbol{n})$ depends on the order in which discovered nodes are encountered.

Future Work and Extensions

Solving cascade dynamics,

- on random graphs containing arbitrary distribution of subgraphs,
- with multiple states (inhibitor, inactive, active, superactive),
- using multiple seeds,
- as a function of time.

Bibliography

[1] J.P. Gleeson, Cascades on correlated and modular random networks, Phys. Rev. E **77**, 046117, (2008).

[2] A. Allard, L. Hébert-Dufresne, P.-A. Noël, V. Marceau, and L.J. Dubé, Exact solution of bond percolation on small arbitrary graphs, EPL **98**, 16001 (2012).

Acknowledgements